Published in Applied Optics, v. 41, 20 August 2002, p. 5,084

Photon-Limited Synthetic Aperture Imaging for Planet Surface Studies

Robert L. Lucke, Code 7218
(lucke@nrl.navy.mil, 202-767-2749)
Lee J Rickard, Code 7210
(rickard@rsd.nrl.navy.mil, 202-404-7877)
U.S. Naval Research Laboratory
Washington DC 20375

Abstract

The carrier-to-noise ratio resulting from phase-sensitive heterodyne detection in a photon-limited synthetic aperture ladar (SAL) is developed, propagated through synthetic aperture signal processing, and combined with speckle to give the signal-to-noise ratio of the resulting image. Carrier- and signal-to-noise ratios are defined in such a way as to be familiar to the optical imaging community. Design equations are presented to show that a 10 μ SAL in orbit around Mars can give centimeter-class resolution with reasonable laser power. SAL is harder to implement in the short-wave infrared and probably not practical at visible wavelengths unless very many separate images can be averaged. Some tutorial information on phase-sensitive heterodyne detection and on synthetic aperture signal processing and image formation is provided.

1. Introduction

Synthetic aperture (SA) imaging has a long history of development in the form of synthetic aperture radar[1,2] (SAR), which has permitted extensive observations of the Earth and of Venus[3]. For Mars, the Moon, and other solar system objects with little or no atmosphere, the same techniques can, in principle, be applied at optical frequencies[4]. A synthetic aperture ladar (SAL) could provide dramatic improvements in either resolution or, compared to SAR, the time needed to record an image, or both. The reduced imaging time results from the shorter time needed by the platform to traverse the SA that produces the same resolution with a shorter wavelength. When the observation range is more than about a hundred kilometers, no other method of imaging can offer centimeter-class resolution with reasonable real aperture size (~ 1 m). SAL can examine a candidate landing zone at the scale of all significant hazards before a vehicle is committed to it, and provide detailed scientific information on small-scale geological features, such as strata or other formations on the walls of river valleys or lava channels. SAL is an active sensing method: it can supply images of places where sunlight is absent, such as winter-time polar regions, and increase coverage in single-pass encounters with asteroids or other objects. A system of this type is probably inappropriate for Earth observations, partly because propagation through the Earth's atmosphere can degrade beam quality substantially at visible and infrared wavelengths, but mostly because it is far easier to achieve high resolution at some point on the Earth by going there than by orbiting a SAL. On other planets, the situation is reversed.

This paper investigates one of the few limits on SAL that is of a theoretical nature: the limit imposed by photon statistics, a limit that is not relevant to SAR. A criterion is developed for the number of photons that are needed from each resolution element of a scene, and design equations are given to evaluate a proposed design with respect to this criterion. The engineering problems of implementing SAL are less easily dealt with. The more prominent are indicated below, but their actual means of solution are resolutely ignored in this paper.

Previous work on SAL (see Green *et al.*[5], Yoshikado *et al.*[6], and references cited therein) has not considered the implications of photon statistics and, in the laboratory, has usually used fixed-frequency CW lasers and measured Doppler shifts from moving targets to create an image. The approach to SAL analyzed here is the SAR technique of transmitting a series of FM-chirped pulses, heterodyning the return signal with a similarly chirped local oscillator (LO), isolating a single range resolution element as a narrow-frequency sub-band of the detector's output (a process called deramping, see Sec. 10.1 of Curlander and McDonough[1] or Sec. 1.3 of Jakowatz *et al.*[2]), and match-filtering data from this sub-band to pick out an azimuth resolution element by its phase history. This technique has recently been demonstrated at 1.55 μ in a laboratory-scale experiment[7], though not yet in the photon-limited regime. This paper examines the effects of photon statistics and of speckle on imagery from a space-based system. We are motivated in part by a desire to bridge the gap between the heterodyne detection and optical imaging communities, so the development will include some relevant tutorial information, but we assume a reasonable degree of familiarity with the physical principles of heterodyne detection ladar (see, for example, Shapiro *et al.*[8] and references cited therein) and SA image formation[1,2]. Park and Shapiro[9] discuss a similar system (their Doppler pulse compression is the equivalent of the phase history matched filter described here), but they emphasize short-range (< 100 km), air-based operation and do not consider photon statistics or speckle. They treat atmospheric propagation extensively, but that topic is largely irrelevant to the applications considered here. Kyle[10] proposes a SAL system that transmits a coded pulse stream, rather than an FM chirp, to resolve range. The method is theoretically sound, but requires very fast modulation of the laser and wide-band detectors. Kyle evaluates his system in much the same way as presented here in Section 6, but drastically overstates the signal-to-noise ratio of his illustrative example. Aleksoff *et al.*[11] show the full potential of SA imaging with a laboratory demonstration of a 3-D SAL, but the method requires 2-D motion of the platform and is therefore unsuitable to the imaging problem considered here.

The system modeled here is a scan-mode SAL that transmits a beam with a ground footprint having an instantaneous diameter that contains M pixels. As the motion of the sensor's platform sweeps the beam along the ground, M pulses, each of time duration T_{pul}, are transmitted during the dwell time, T_{dw}, the time a single pixel remains illuminated. With pulse length T_{pul}, the minimum detectable frequency difference in the heterodyne signal is $\delta f = 1/T_{pul}$, and this, combined with the chirp rate, determines the minimum resolvable range element. Azimuthal SA processing requires measuring both the amplitude and phase of the light scattered from the scene, and at optical frequencies this can be done only with heterodyne detection. The fact that phase must be measured separates SAL from conventional optical heterodyne systems, which are used as sensitive detectors of narrow-band light, but measure only the number of photons received, not their phase. For this reason, SAL necessitates a more thorough treatment of shot noise than is normally required.

For a photon-limited direct detector, the number of signal photons detected in a single measurement is known, but the same cannot be said for a photon-limited heterodyne detector. Because of shot noise from the LO, it is impossible to conclude that a particular number of signal photons was detected in a measurement. Consequently, the value of *n*, the number of signal photons inferred from the heterodyne measurement, is not restricted to integral values and is treated as a continuous variable when its probability density function (PDF) is considered. The PDF is needed to calculate the carrier-to-noise ratio (CNR) and, combined with speckle, the signal-to-noise ratio (SNR). CNR is an unfamiliar term in normal, direct-detection, optical imaging: it means SNR before the effect of speckle is included. The definition of CNR normally used for heterodyne detection is a legacy of its RF origin and leads to a photon-limited CNR proportional to the number of signal photons instead of the square root of this number. The photon-limited CNR for SAL will be defined to be proportional to the square root of the number of signal photons, a definition more familiar to the optical imaging community. CNR and SNR for SAL will be compared to those for a direct detection system that detects the same number of photons from one polarization of the light returned from a coherently illuminated scene (recall that scattering from ordinary surfaces randomizes polarization). In other words, SAL will be compared to a direct-detection system with a polarizer in it. CNRs and SNRs can always be improved by a factor of $\sqrt{2}$ by measuring both polarizations, but this is far easier to do with a direct-detection system (just remove the polarizer!) than with a heterodyne system, for which a beamsplitter and an additional detection channel must be added. Speckle limits the SNR of single-look imagery to, at most, unity for SAL, just as it does for SAR or for direct detection. Fortunately, the limit can be closely approached when only a few photons per pixel are received.

Section 2 describes phase-sensitive heterodyne detection, with emphasis on the fact that signal and noise are complex numbers in Fourier space. Section 3 derives the appropriate CNR for an imaging system and compares it to the traditional RF definition, Section 4 propagates signal and noise through synthetic aperture processing, and Section 5 combines the result with speckle to produce the SNR of the SAL image. Section 6 presents design equations, with emphasis on the specifications of the laser, and candidate designs for a Mars-orbiting SAL. Section 7 summarizes some of the applications and limitations of SAL.

2. Phase-Sensitive Heterodyne Detection

A light wave with frequency f and phase ϕ is described by $E\exp(2\pi ift + i\phi)$ with E real and non-negative, and the units of E are chosen so that power is related to the electric field by

$$P = \int_{area} \frac{1}{2}\left|E\exp(2\pi ift + i\phi)\right|^2 dxdy \equiv \frac{1}{2}A_d\,\overline{E}^2 \quad , \tag{1}$$

where the integral is over the area, A_d, of the detector, E may be, and usually is, a function of position on the detector, and $\overline{E}$ is the appropriate average. To relate $\overline{E}$ to N, the average number of photons in pulse time T_{pul} (N need not be an integer), we write $P = h\nu N/T_{pul}$, where $h = 6.63\times10^{-34}$ joule-sec and ν is the frequency of the light, to find that

$$\overline{E} = \sqrt{\frac{2h\nu N}{A_d T_{pul}}} \quad . \tag{2}$$

The frequency of an FM-chirped waveform is $f = f_0 + \dot{f}t$, where $\dot{f} = df/dt$ is the chirp rate, so the field of the LO is $E_L \exp[2\pi i(f_0 + \tfrac{1}{2}\dot{f}t)t]$. The field of the signal from a single range resolution element, being displaced in time by some amount Δt and having an arbitrary phase ϕ_{S0} with respect to the LO, is $E_S \exp\{2\pi i[f_0 + \tfrac{1}{2}\dot{f}(t + \Delta t)](t + \Delta t) + i\phi_{S0}\}$. In heterodyne detection the fields are combined on the detector to yield a detector output current given by

$$\begin{aligned} I_d &= \eta_d \frac{q_e}{h\nu} \int_{area} \frac{1}{2} \left| E_L \exp[2\pi i(f_0 + \tfrac{1}{2}\dot{f}t)t] + E_S \exp\{2\pi i[f_0 + \tfrac{1}{2}\dot{f}(t + \Delta t)](t + \Delta t) + i\phi_{S0}\} \right|^2 dxdy \\ &= \eta_d \frac{q_e}{h\nu} \left[\frac{1}{2} A_d \overline{E}_L^2 + \frac{1}{2} A_d \overline{E}_S^2 + A_d \sqrt{\eta_h}\, \overline{E}_L \overline{E}_S \cos(2\pi \Delta f t + \phi_S) \right] \\ &= q_e \eta_d \frac{N_L + N_S}{T_{pul}} + 2 q_e \eta_d \frac{\sqrt{\eta_h N_L N_S}}{T_{pul}} \cos(2\pi \Delta f t + \phi_S) \ , \end{aligned} \tag{3}$$

where η_d is the detector's quantum efficiency, assumed constant across the detector, q_e is the charge of an electron, $q_e/h\nu$ performs the detector's transducer function of replacing $h\nu$ by q_e, η_h is the heterodyne mixing efficiency[12], $\Delta f = \dot{f}\Delta t$ is the beat frequency, $\phi_S = \phi_{S0} + 2\pi(f_0 \Delta t + \tfrac{1}{2}\dot{f}\Delta t^2)$, and Eq. (2) has been used. The first term in the third equality of Eq. (3), when multiplied by T_{pul}/q_e, is the total number of electrons generated (= photons detected). The second term identifies the range element in question by its beat frequency Δf. A different range element yields a different Δf, a relation that will be stated precisely in Section 6. Eq. (3) is most easily understood from the point of view of the semiclassical theory[13], that the field itself may be treated classically, that is, without intrinsic fluctuations. Fluctuations in the number of photons detected results from a stochastic interaction between the electromagnetic field and the detector: shot noise, which is treated below.

Eq. (3) is written for a single range resolution element. In the detector's actual output, there are M such terms, having M different frequencies, one for each range resolution element in the footprint. In order to satisfy the Nyquist criterion, the detector's output is digitized with (at least) $2M$ samples over the time T_{pul}, and the value of the Δf component of the discrete Fourier transform (DFT) of these samples is

$$\begin{aligned} D(\Delta f) &= \frac{T_{pul}}{2M} \sum_{m=0}^{2M-1} 2 q_e \eta_d \frac{\sqrt{\eta_h N_L N_S}}{T_{pul}} \cos(2\pi \Delta f t_m + \phi_S) \exp(-2\pi i \Delta f t_m) \\ &= q_e \eta_d \sqrt{\eta_h N_L N_S}\, \exp(i\phi_S) \quad , \end{aligned} \tag{4}$$

where $T_{pul}/2M$ normalizes the DFT so that its DC component is the total charge generated and $t_m = mT_{pul}/2M$ is the time of the mth sample. $D(\Delta f)$ is divided by $q_e \eta_d (\eta_h N_L)^{1/2}$ to obtain the desired value, $D'(\Delta f) = N_S^{1/2} \exp(i\phi_S)$, that is needed for SA processing. It is a basic property of the DFT that the separation between the DFT's discrete frequency components is $\delta f = 1/T_{pul}$, so D' is the signal over bandwidth δf (*i.e.*, from one range resolution element) at a frequency displaced by Δf from the frequency of the LO.

In the photon-limited regime, the dominant source of noise is shot noise from the total number of photons detected, which is $\eta_d(N_L + N_S)$. Normally $N_L >> N_S$, and that approximation will be used here. $N_L >> 1$ always. As shown in Appendix A, the noise at any frequency is described by a 2-D Gaussian distribution [Eq. (A1) with $s = 0$] with, replacing N in Eq. (A5) by $N_L + N_S$, $\sigma^2 = q_e^2\eta_d(N_L + N_S)/2 \approx q_e^2\eta_d N_L/2$. This is the noise on the signal D. If a random variable is divided by a constant to obtain a new random variable, the variance of the old variable must be divided by the square of the constant to obtain the variance of the new one. Since D is divided by $q_e\eta_d(\eta_h N_L)^{1/2}$ to obtain the desired value, D', σ^2 must be divided by the square of this factor, $(q_e\eta_d)^2\eta_h N_L$, to obtain

$$\sigma'^2 = \frac{1}{2\eta_d\eta_h} \quad . \tag{5}$$

That is, the PDF of the random variable D' plus noise is a 2-D Gaussian centered on $N_S^{1/2}\exp(i\phi_S)$ with width given by σ'^2, as illustrated schematically in Figure 1. With signal and noise now specified, we are ready to describe SA processing and see how noise propagates through it, but it is instructive to pause at this point to examine the CNR and the number and phase uncertainties of heterodyne detection.

3. Carrier-to-Noise Ratio

As stated above, CNR is SNR before speckle is taken into account, so the results of this section apply to a coherent-light sensor that makes repeated measurements without changing the part of the speckle field it samples. In optical imaging, the normal definition of SNR or CNR is the ratio of the magnitude of a signal to the standard deviation (square root of the variance) of the signal's estimator. For photon-counting direct detection, the number of detected photons, n, follows Poisson statistics with $\langle n\rangle = \eta_d N_S$. Now, n must be divided by η_d to obtain an estimate of the signal: $\langle n/\eta_d\rangle = N_S$. For the Poisson distribution the variance is equal to the mean, that is, $\mathrm{Var}(n) = \langle n\rangle = \eta_d N_S$, which must be divided by η_d^2 to obtain N_S/η_d, the variance of the estimator of the signal. Thus CNR $= N_S/(N_S/\eta_d)^{1/2} = (\eta_d N_S)^{1/2}$, as expected.

For heterodyne detection, the result of measuring the return from a single pulse is a complex number, $r\exp(i\phi)$, equal to D' plus noise, from which an estimate of D' must be derived. As shown in Fig. 1, $r\exp(i\phi)$ is distributed according to Eq. (A1) with $s = N_S^{1/2}$ (without loss of generality, we have set $\phi_S = 0$) and $\sigma = \sigma'$ from Eq. (5). The magnitude, r, is the square root of the number of photons inferred from the measurement: $r = n^{1/2}$. Now $\langle n\rangle = \langle r^2\rangle = \langle x^2\rangle + \langle y^2\rangle$, and the Gaussian moments of Eq. (A1) are easily evaluated to show that $\langle n\rangle = s^2 + 2\sigma'^2$. It is only slightly less easy to use $\langle n^2\rangle = \langle r^4\rangle = \langle (x^2 + y^2)^2\rangle$ to show that $\langle n^2\rangle = s^4 + 8s^2\sigma'^2 + 8\sigma'^4$ and therefore that the variance of the 1-D distribution of n is

$$\begin{aligned} \mathrm{Var}(n) &= \left\langle n^2\right\rangle - \left\langle n\right\rangle^2 = 4s^2\sigma'^2 + 4\sigma'^4 \\ &= 4N_S\sigma'^2 + 4\sigma'^4 \quad . \end{aligned} \tag{6}$$

An unbiased estimator of N_S is $n - 2\sigma'^2$, since $\langle n - 2\sigma'^2 \rangle = s^2 = N_S$. Since, from Eq. (5), σ'^2 = constant, the variance of this estimator is the same as the variance of n. Using the definition given above, the CNR of heterodyne detection for imaging applications is the ratio of N_S to the standard deviation of its estimator:

$$\begin{aligned} \mathrm{CNR}_{\mathrm{IM}} = \frac{N_S}{\sqrt{\dfrac{2N_S}{\eta_d \eta_h} + \dfrac{1}{\eta_d^2 \eta_h^2}}} &\approx \sqrt{\frac{\eta_d \eta_h N_S}{2}} \quad \text{for } N_S \;?\; 1/(\eta_d \eta_h) \\ &\approx \eta_d \eta_h N_S \quad \text{for } N_S = 1/(\eta_d \eta_h) \;, \end{aligned} \tag{7}$$

where Eqs. (5) and (6) have been used. Taking $\eta_d = \eta_h = 1$ in the first approximation shows that the best possible CNR of heterodyne detection is a factor of $\sqrt{2}$ below the best possible CNR of direct detection. For $N_S << 1/(\eta_d \eta_h)$, $\mathrm{CNR}_{\mathrm{IM}}$ is proportional to the number of photons detected, rather than to the square root of this number, a fact that will be revisited in Section 5 [see discussion below Eqs. (13) and (14)] where it is found to apply also to the SNR of SA imaging.

Heterodyne detection was first done in RF work, where CNR is defined as the ratio of signal power to noise power, where signal power means the square of the value of the heterodyne-detected signal, which is $(N_S^{1/2})^2 = N_S$, and noise power means $2\sigma'^2$. Thus

$$\mathrm{CNR}_{\mathrm{RF}} = \frac{N_S}{2\sigma'^2} = \eta_d \eta_h N_S = 2\left(\mathrm{CNR}_{\mathrm{IM1}}\right)^2 \quad , \tag{8}$$

where $\mathrm{CNR}_{\mathrm{IM1}}$ is taken from the first approximation in Eq. (7). The first equality in Eq. (8) matches, for two examples, Eq. (12) of Park and Shapiro[9] and Eq. (1) of Shapiro[14], once the conversion to photons has been made. If the RF expression for CNR is used to discuss an optical imaging system, large overstatements of CNR can result, compared to what is normally expected in optical imaging.

In the ideal case, $\eta_d = \eta_h = 1$ and the variance is $2\sigma_o'^2 = 1$. The magnitude of D' is the square root of a (not necessarily integral) number of photons. The units of σ'^2 may therefore be regarded as photons, so that $2\sigma_o'^2 = 1$ photon. This is the variance (in Fourier space) for a measurement done over time interval T_{pul} and frequency interval $\delta f = 1/T_{pul}$, so $2\sigma_o'^2$ may be generalized to $2\sigma_o'^2 = [1 \text{ photon/(secHz)}] \times T_{pul}\delta f$, since the time-bandwidth product of the measurement is $T_{pul}\delta f = 1$. This generalization illustrates the oft-heard statement that heterodyne detection adds noise at the rate of 1 photon/(secHz). But this statement can be misleading to those accustomed to photon-limited imaging because $2\sigma_o'^2$ is a variance (= noise power), not a standard deviation, and it is the variance of the 2-D distribution of D' plus noise, shown in Fig. 1, not the variance of the 1-D distribution of the number of photons, which is Var(n), given in Eq. (6). Another way to express this variance is to multiply it by $h\nu$ and write it in terms of power as $2\sigma_o'^2 = h\nu$ wt/Hz, or, more generally, as $2\sigma'^2 = h\nu/(\eta_d \eta_h)$ wt/Hz. This form, multiplied by a receiver bandwidth, appears as the denominator of the expressions for CNR given by Park and Shapiro[9] and Shapiro[14].

The 1-D distribution of the number of photons can be found from Goodman's[15] Eq. (2.9-20) or (2.9-27). These equations give, respectively, exact and approximate forms of the PDF of $n^{½}$ as, in Goodman's notation, $P_A(a)$, where $a = n^{½}$. Using $n = a^2$, the PDF transformation method in Goodman's Sec. 2.5.2 shows that the distribution of the number of photons inferred from a heterodyne detection of N_S is $P_{N_S}(n) = P_A(a)/(2a)$. This distribution is stated as Eq. (19) of Shapiro and Wagner[16], but is not given here because all we need (see Section 5) are its first and second moments, $\langle n \rangle$ and $\langle n^2 \rangle$, which have been given above Eq. (6). Since we have not found them in any reference, we exhibit, for small N_S, the uncertainties, Δn and $\Delta\phi$, due to heterodyne measurement of the values of n and ϕ, from which N_S and ϕ_S are estimated. The term uncertainty is used in place of standard deviation because that is the usual terminology of quantum mechanics. We already know from Eq. (6) or the denominator of the first equality in Eq. (7) that $\eta_d\eta_h\Delta n = (2\eta_d\eta_h N_S + 1)^{½}$. This way of expressing Δn is chosen because it allows a single curve to show Δn for all values of $\eta_d\eta_h$. We again set $\phi_S = 0$ so $\Delta\phi = \langle\phi^2\rangle^{½}$, which can be evaluated by numerical integration using the PDF in Eq. (A1) and $\phi = \tan^{-1}(y/x)$, or using Goodman's Eq. (2.9-25). The result is plotted in Figure 2, along with $\Delta\sin\phi$ and the products $\Delta n\Delta\phi$ and $\Delta n\Delta\sin\phi$. $\Delta\sin\phi$ is included because $\sin\phi$ rather than ϕ is the true quantum mechanical observable[17]. For $\eta_d = \eta_h = 1$, $\Delta n\Delta\phi \geq 1$ for heterodyne detection, as also found by, for example, Shapiro and Wagner[16].

4. Synthetic Aperture Processing

Sec. 2 described the measurement of D' plus noise, where D' is the complex amplitude of the wave reflected from one range resolution element and σ' describes the noise of the measurement. Figure 3 shows a range resolution element, divided into M pixels. The pixel labeled $m = 0$ has just entered the illuminated region. Each illuminated pixel contributes a phasor, $A_m = a_m\exp(i\alpha_m)$, with amplitude a_m and intrinsic phase α_m, to D'. The intrinsic phase depends on the detailed structure of the element and on the viewing geometry and may be regarded as random. Thus the average number of photons contributing to D' is the result of the sum of the random phasors A_m, and the discussion in the first paragraph of Appendix A shows that $a_{m,\mathrm{rms}} = (N_S/M)^{½}$ in order that the A_m add up to a complex number with magnitude $N_S^{½}$.

The phase of each of the A_m is modified by the curvature of the wave front as indicated in Fig. 3. Referenced to zero at the center of the wave front ($m = M/2$), the light described by phasor A_m must travel a greater distance by the amount $\Delta l_m = [(m - M/2)p]^2/R$, where p is the size of a pixel and R is the range, to return to the detector. A_m is therefore multiplied by the complex phase coefficient $C_m = \exp(i\Delta l_m \times 2\pi/\lambda)$, and the sum of all these contributions makes up the signal: $D' = \Sigma A_m C_m$. Fig. 3 and the expression used for Δl_m assume that the beam direction is exactly perpendicular to the velocity vector of the transmitter. Relaxing this assumption changes the expression for the Δl_m, but not the results of the analysis. The noise, or error, denoted by E_m, adds to the signal to produce the result of a single measurement, $\Sigma A_m C_m + E_m$, as indicated on the following page by the first row of Table 1 (less the last column, which indicates multiplication by C_m^*). The second row shows the contributions to the measurement of the second pulse, when the illuminated region has moved by one pixel, and so on, until the $M - 1$

row describes the pulse having the last contribution from the $m = 0$ pixel. Synthetic aperture processing applies a matched filter to pick out the phase history of a particular pixel as it passes through the beam's footprint. As indicated in Table 1, and described further below, the filter picks out A_0. Appendix B gives mathematical detail and shows how the high resolution characteristic of SA processing is obtained. The C_m* are indexed upward by one row to pick out A_1, downward to pick out A_{-1}, *etc*.

{			A_0C_0	$+A_1C_1+\ldots+A_{M-2}C_{M-2}$		$+A_{M-1}C_{M-1}$	$+E_0$ }	$\times C_0^*$
{		$A_{-1}C_0$	$+A_0C_1$	$+A_1C_2+\ldots+A_{M-2}C_{M-1}$			$+E_1$ }	$\times C_1^*$
{	$A_{-2}C_0$	$+A_{-1}C_1$	$+A_0C_2$	.	.		$+E_2$ }	$\times C_2^*$
	.	.	.	.	.		.	.
	.	.	.	.			.	.
	.	.	.	.			.	.
{. . .	$A_{-2}C_{M-3}+$	$A_{-1}C_{M-2}+$	A_0C_{M-1}				$+E_{M-1}$}	$\times C_{M-1}^*$
{. . .	≈ 0	$+\approx 0$	$+MA_0$	$+\approx 0$	$\ldots+\approx 0$	$+\approx 0$	$+G$ }	

Table 1. Contributions of M pixels to a single measurement are multiplied by the coefficients C_m* and added (vertically) to yield the output of the phase-history matched filter.

For clarity of presentation in Table 1, the profile of the illuminating beam is treated as uniform, when in reality it would have a Gaussian or perhaps an Airy shape. A more careful treatment includes the beam's non-uniform profile in the matched filter, but does not change the final results here or in Appendix B.

The next-to-last column in Table 1 contains the E_m, which are the error contributions from shot noise. The E_m are random numbers distributed according to Eq. (A1) with $s = 0$ and $\sigma = \sigma'$ from Eq. (5). The last column shows the coefficients of the phase history matched filter. These coefficients are calculated from the known geometry of the observation. The bottom row, which is the sum of the rows above it after the multiplications by C_m* have been done, shows the contributions to the output of this filter. The output of the filter is comprised (mostly) of the phasor MA_0 plus noise given by $G = \Sigma E_m C_m$*. The result of multiplying the random complex number E_m by the unit-magnitude phasor C_m* is again a random complex number, so G, the sum of M such numbers, is a Gaussian-distributed random complex number with $\sigma''^2 = M\sigma'^2$. The total number of photons from the $m = 0$ pixel in the final image is denoted N_p, so N_p/M are contributed by each pulse, that is, $a_0^2 = N_p/M$. Note that a_0^2 may be small compared to unity. The PDF of $MA_0 + G$ is centered on $Ma_0 = (MN_p)^{1/2}$, while the desired signal is N_p. Therefore, the output of the matched filter must be divided by $M^{1/2}$ to obtain a PDF centered on $N_p^{1/2}$. Accordingly, σ''^2 must be divided by M, returning us to σ'^2 as before. Thus, the error in the output of the matched filter is the same as the error in the measurement of $N_S^{1/2}$, so the PDF for N_p is the same as that given for N_S at the end of Sec. 3, with N_S replaced by N_p. That is, it is $P_{N_p}(n)$, and Var(n) and the CNR for N_p are given by Eqs. (6) and (7) with N_S replaced by N_p.

The average number (averaged over speckle, which is treated in the next section) of photons received per pulse is $N_{S,\mathrm{ave}}$ and the average number of photons per pixel in the final image is $N_{p,\mathrm{ave}}$. Since the photons received from M pulses are distributed over M pixels, $N_{S,\mathrm{ave}} = N_{p,\mathrm{ave}}$, that is, the average number of photons inferred from a single heterodyne detection is the same as the average number in each pixel in the final image.

It was stated above that the C_m* are calculated from the geometry of the observation. This calculation is a harder problem for SAL than for SAR because the accuracy of the calculation is set by the wavelength of the radiation, which is typically about four orders of magnitude smaller for SAL. (Fortunately, the range need not be known to wavelength accuracy: referring to Fig. 3, it is easy to show that the curvature of a wave front at 999 km range is insignificantly different from its curvature at 1,000 km for the small footprints possible with SAL – see Section 6.) Platform vibrations represent random optical path length errors that must be either eliminated or measured and compensated (by adjusting the measured value of ϕ_S) if they are an appreciable fraction of a wavelength. If this cannot be done, the phase of one returned pulse will differ from that of another by an unknown amount and the output of the matched filter will not be the desired quantity. The problem is alleviated by the fact that maintaining coherence for the duration of a pulse, which is only $1/M$ of the whole dwell time, is all that is essential to SA processing. If the phase changes sufficiently slowly from pulse to pulse, it can be corrected in data processing, *e.g.*, by trying different phases for each pulse until the phases that produce an image are found. This technique, called "focusing", is used for a variety of reasons in SAR and is feasible as long as the phase changes are slow[2].

5. Speckle and SNR

The most severe limit on an imaging system that uses coherent light is speckle. In the foregoing, the variance due to shot noise of a single pixel in the final SA image has been calculated. But if that pixel were viewed from a different direction, the sensor would be in a different part of the pixel's speckle field and a different value of N_p would be observed. Or if there are many pixels in the scene with the same reflection properties, they will produce different values of N_p because different parts of their speckle fields are sampled. To find the resulting SNR in an image, the speckle and shot noise contributions to variance must be combined.

With $s = 0$, Eq. (A1) describes the phasor distribution of speckle. The resulting intensity PDF [see Goodman[15], Eq. (7.5-1)] is,

$$P_S\left(N_p\right) = \frac{1}{N_0}\exp\left(-\frac{N_p}{N_0}\right) \quad , \tag{9}$$

where $N_0 = \langle N_p \rangle$ is the average of N_p taken over many realizations of speckle. The second moment of this distribution is $\langle {N_p}^2 \rangle = 2{N_0}^2$, so the variance is $\mathrm{Var}(N_p) = {N_0}^2$. This leads to the familiar result that the SNR due to speckle for a single polarization is unity: $\mathrm{SNR} = N_0/[\mathrm{Var}(N_p)]^{½} = 1$. Eq. (9) is given as a continuous function. In high-intensity speckle this is justified because $N_0 >> 1$, in our case it is justified because both N_p and N_0 are averages and are normally not integers.

For a fixed realization of speckle, the variation in the number of photons, n, that contribute to a particular pixel in a single image is due only to shot noise from the heterodyne detection. As stated in Sec. 4, the PDF of this 1-D distribution, $P_{N_p}(n)$, was introduced [as $P_{N_S}(n)$] at the end of Sec. 3. We have not calculated $P_{N_p}(n)$ explicitly, but, using $s^2 = N_p$ in the expressions for $\langle n \rangle$ and $\langle n^2 \rangle$ given above Eq. (6), have calculated its first moment to be $N_p + 2\sigma'^2$ and its second to be ${N_p}^2 + 8N_p\sigma'^2 + 8\sigma'^4$. Combining the distribution of speckle realizations given by Eq. (9) with $P_{N_p}(n)$ gives the total probability of getting a particular value n:

$$P_T(n) = \int_0^\infty P_{N_p}(n) P_S\left(N_p\right) dN_p \quad , \tag{10}$$

which may be stated in words as the probability of measuring the value n given N_p, summed over the probability of N_p. The moments of n are

$$\begin{aligned} \langle n \rangle &= \int_0^\infty n P_T(n)\, dn = \int_0^\infty \int_0^\infty n P_{N_p}(n)\, dn\, P_S\left(N_p\right) dN_p \\ &= \int_0^\infty \left(N_p + 2\sigma'^2\right) P_S\left(N_p\right) dN_p \\ &= N_0 + 2\sigma'^2 \quad , \end{aligned} \tag{11}$$

as expected, and

$$\begin{aligned} \left\langle n^2 \right\rangle &= \int_0^\infty \int_0^\infty n^2 P_{N_p}(n)\, dn\, P_S\left(N_p\right) dN_p \\ &= \int_0^\infty \left(N_p^2 + 8N_p\sigma'^2 + 8\sigma'^4\right) P_S\left(N_p\right) dN_p \\ &= 2N_0^2 + 8N_0\sigma'^2 + 8\sigma'^4 = 2(N_0 + 2\sigma'^2)^2 \quad , \end{aligned} \tag{12}$$

so $\mathrm{Var}(n) = (N_0 + 2\sigma'^2)^2$, and, using σ'^2 from Eq. (5), the SNR for photon-limited SA imaging is

$$\begin{aligned} \mathrm{SNR}_{\mathrm{SA}} = \frac{N_0}{N_0 + \dfrac{1}{\eta_d \eta_h}} &\approx 1 && \text{for } N_0 \,?\, 1/(\eta_d \eta_h) \,, \\ &\approx \eta_d \eta_h N_0 && \text{for } N_0 = 1/(\eta_d \eta_h) \,, \\ &= 1/2 && \text{for } N_0 = 1/(\eta_d \eta_h) \,. \end{aligned} \tag{13}$$

These steps can be repeated for direct detection: in Eqs. (11) and (12), $P_{N_p}(n)$ is replaced by the Poisson distribution having mean $\eta_d N_p$ and the integral over n by the discrete sum over n

appropriate for a Poisson distribution. The results are $\langle n/\eta_d \rangle = N_0$, $\langle (n/\eta_d)^2 \rangle = 2N_0^2 + N_0/\eta_d$, and $\mathrm{Var}(n/\eta_d) = N_0^2 + N_0/\eta_d$, so the SNR for photon-limited direct detection (DD) of one polarization of coherent light is

$$\begin{aligned} \mathrm{SNR}_{\mathrm{DD}} = \sqrt{\frac{N_0}{N_0 + \frac{1}{\eta_d}}} \quad &\approx 1 \qquad \text{for } N_0 \,?\; 1/\eta_d \ , \\ &\approx \sqrt{\eta_d N_0} \quad \text{for } N_0 = 1/\eta_d \ , \\ &= \sqrt{\frac{N_0}{N_0 + 1}} \quad \text{for } \eta_d = 1 \ , \end{aligned} \tag{14}$$

which, for $\eta_d = 1$, is the same as Goodman's[15] Eq. (9.2-18) once the identification $\overline{K} = N_0$ has been made.

Eqs. (13) and (14) show the saturation effect[8] expected when speckle is the dominant source of noise: when $N_0 >> 1/(\eta_d \eta_h)$, SNR ≈ 1 in both cases and higher values of N_0 do not improve SNR. When $N_0 = 1/(\eta_d \eta_h)$, the SNR of SAL imagery is ½ – only a factor of two below the limiting value of 1. SNR can, of course, be improved at the cost of complexity by measuring both polarizations, at the cost of resolution by combining pixels in one image, as is often done for SAR, and/or at the cost of more observation time by combining images that sample different parts of the speckle field.

For small N_0, $\mathrm{SNR}_{\mathrm{SA}}$ is proportional to the number of photons detected, while $\mathrm{SNR}_{\mathrm{DD}}$ has the more familiar property of being proportional to the square root of this number. Now, if multiple images from uncorrelated parts of the speckle field are added together, both $\mathrm{SNR}_{\mathrm{SA}}$ and $\mathrm{SNR}_{\mathrm{DD}}$ improve only as the square root of the number of images combined (this is the normal statistical expectation, see also p. 217 of Curlander and McDonough[1]). This means that it is much harder to make up for a low count rate by adding SA images, compared to DD images, as the following numerical example shows. Taking $\eta_d = \eta_h = 1$ for simplicity, $N_0 = 0.1$ implies $\mathrm{SNR}_{\mathrm{SA}} \approx 0.1$ and $\mathrm{SNR}_{\mathrm{DD}} \approx 0.3$. It takes 9 images to improve $\mathrm{SNR}_{\mathrm{DD}}$ to 0.9, but 81 to improve $\mathrm{SNR}_{\mathrm{SA}}$ to 0.9. This shows the importance of designing an SA system to meet the criterion $N_0 = 1/(\eta_d \eta_h)$ implied by Eq. (13). If this condition is not met, very many single-look images will have to be combined just to approach an SNR of unity.

The astute reader may ask about taking the multiple images while maintaining phase coherence across a larger part of the speckle field, *i.e.*, using a larger SA. But with, say, twice the SA, SA processing yields twice as many pixels, each having half the extent in the azimuth direction as the original and each receiving the same number of photons as the original pixel. Adding these pixels together to match the original pixel improves $\mathrm{SNR}_{\mathrm{SA}}$ by $\sqrt{2}$, the same result as adding two successive, separately processed images, so the improvement is the same whether the multiple images are taken coherently or incoherently. It may be helpful to remind the reader that, in SA imaging, the length of the SA is also the correlation length of the speckle field of a pixel.

6. Space-Based SAL Design Equations

In Sec. 5, we showed that the total number of photons present at the detector from one pixel in the scene needs to be $N_0 \approx 1/(\eta_d \eta_h)$ for a worthwhile SNR in an image (multiple images can then be added to improve SNR in the usual way). In this section, we calculate the number of photons per pixel that would be received by a Mars-orbiting SAL. The result is easily extended to other solar system bodies. Parameters are given below for a baseline system using 2 μ light and for two variations using 10 μ light. The baseline system assumes a beam footprint of 10 m and a resolution of 0.1 m. The 10-m footprint exposes a basic limitation of SAL – supplying enough photons to cover a substantially larger footprint requires prohibitive laser power. A range of 1,000 km is chosen in part to give an easily-scaled parameter, in part because it is about the maximum range at which SAL is likely to be feasible. The footprint size is assumed to be determined by the diffraction limit of the transmitting aperture. Circular apertures for the transmit and receive optics are assumed, but the area ratio of circles to squares ($\pi/4$) is ignored. Also ignored is the difference between the resolution measured perpendicular to the beam and measured on the ground. With the resolution measured perpendicular to the beam, the pixels are assumed square, that is, when the ground surface is inclined at 45° with respect to the beam, the laser is assumed to have sufficient chirp capability to yield the same cross-track resolution as SA processing yields in the azimuth (along-track) direction.

A degree of flexibility that is essential to an effective SAL is added by assuming that the beam can be rapidly repositioned, either by steering the beam or rotating the satellite, so that a footprint can be scanned more than once and/or neighboring footprints can be covered. This capability is needed to make up for the low SNR and small footprint inherent to SAL. Beam repositioning allows improvements in SNR by taking images from different parts of the speckle field, and in area coverage by taking a mosaic of images. Further, we assume that coherence can be maintained for up to N_{sc} scans of the same footprint, so that multi-scanning can also improve resolution. $N_{sc} = 1$ for normal scan mode operation and in this mode SA processing is contained in the assumption that the pixel size, p, is one-half the diameter of the transmitting aperture, D_T, as shown in Eq. (B5). Scanning the footprint N_{sc} times means that the platform traverses an N_{sc}-times-longer synthetic aperture and therefore that the resolution of the image can be N_{sc} times better, *i.e.*, that the pixel size is given by $p = D_T/(2N_{sc})$. If N_{sc} is allowed to become large, this process approaches spotlight-mode SA imaging[2], a subject that is not considered here. The total dwell time, T_{dw}, is defined to be the time that a single point on the ground is illuminated by the beam. This is the time it takes to move the beam the length of the footprint at the speed of the orbiting platform, multiplied by N_{sc} to account for multiple scans.

SAR systems normally use the same antenna for transmission and reception, but this is not essential for SA imaging. For SAL, the receiver will be assumed to have a different aperture, with K times larger diameter than the transmitter. The K-times-larger aperture collects K^2 as much light from a ground pixel and has a K-times-smaller footprint than the transmitter. There must therefore be K^2 heterodyne detectors in the focal plane of the receiver instead of one, and light from the first M/K pixels shown in Fig. 3 is detected by one of these, light from the next M/K pixels by another, and so on. Thus, the phase history of a pixel indicated in Table 1 must be traced through the outputs of K detectors. There is no problem with this in principle, "only" in engineering.

Basic design parameters (general terms and a baseline numerical example are given)

1) $\lambda = 2\ \mu = 2\times10^{-6}$ m
2) P = laser output power = 1 kilowatt (time-averaged)
3) transmit optics area = $D_T^{\,2} = (0.2\text{ m})^{\,2}$
4) receive optics area = $D_R^{\,2} = K^2\,D_T^{\,2} = (1\text{ m})^{\,2}$ ($\Rightarrow K = 5$)
5) R = range to scene = 1,000 km = 10^6 m
6) V = platform speed = 3.5×10^3 m/sec for low-altitude Mars orbit
7) ρ = surface reflectance = 0.1 (changes with wavelength), Lambertian distribution
8) η_{op} = combined efficiency of transmit and receive optics = 0.5
9) N_{sc} = number of scans of footprint = 1

Derived quantities

1) F = footprint size = $(\lambda/D_T)R$ = 10 m ($\Rightarrow D_T = \lambda R/F$)
2) p = pixel size = $D_T/(2N_{sc})$ = 0.1 m (from SA processing)
3) Ω = collection solid angle = $D_T^{\,2}/R^2 = [(1\text{ m})/(10^6\text{ m})]^2 = 10^{-12}$ steradians
4) T_{dw} = dwell time = ($N_{sc}\times$footprint length)/(platform speed) = $N_{sc}F/V$ = 2.9 msec
5) M = number of pulses that illuminate one pixel = F/p = 100
6) T_{pul} = pulse time $\le T_{dw}/M = N_{sc}p/V$ = 29 μsec

Derived quantities 5 and 6 are determined by the SA processing requirement that there be one transmitted pulse per azimuthal resolution element. The pulse repetition frequency (PRF) is

$$\text{PRF} = \frac{V}{N_{sc}\,p} = 3.5\times10^4\left(\frac{1}{N_{sc}}\right)\left(\frac{0.1\,\text{m}}{p}\right)\left(\frac{V}{3.5\,\text{km/sec}}\right)\ \text{pulses/sec} \quad . \tag{15}$$

We saw in Eq. (4) that the Nyquist criterion requires that each pulse be sampled at least $2M$ times to recover M range resolution elements, so the data sampling rate, SR, must be $\text{SR} \ge 2M/T_{pul} \ge 2FV/(N_{sc}p^2)$ = 7 MHz for the numerical example, a modest requirement.

To see what the frequency range of the laser's chirp must be, we first observe that the two-way transit time of a wave front across a range increment δl is $\delta t = 2\delta l/c$. In Eq. (3), we saw that if the laser's frequency is varied linearly through a total chirp range Δf_{ch} in the pulse time T_{pul} ($f = \Delta f_{ch}/T_{pul}$), the change in beat frequency caused by the time increment δt is $\delta f = 2(\Delta f_{ch}/T_{pul})(\delta l/c)$. We have already seen that the minimum detectable frequency difference is $\delta f = 1/T_{pul}$, so setting $\delta l = p$ (if the surface is inclined at 45°, a range resolution of p implies an image resolution of p measured perpendicular to the beam and $\sqrt{2}\,p$ measured along the surface) gives

$$\Delta f_{ch} = \frac{c}{2p} = 1.5\left(\frac{0.1\,\text{m}}{p}\right)\ \text{GHz} \quad . \tag{16}$$

The length of the synthetic aperture is the distance traversed by the platform in the dwell time. It can be written in a number of useful forms, some of which are

$$L_{\mathrm{SA}} = VT_{dw} = N_{sc}F = \frac{\lambda R}{2p} \quad . \tag{17}$$

The diameter of the transmitter needed to give the desired footprint and the consequent pixel size that results from SA processing are related by

$$D_T = 2N_{sc}\,p = \frac{\lambda R}{F} = 0.2\left(\frac{\lambda}{2\,\mu}\right)\left(\frac{10\ \mathrm{m}}{F}\right)\left(\frac{R}{1{,}000\,\mathrm{km}}\right) \text{ meters.} \tag{18}$$

The time-averaged laser power within the usable footprint is taken to be $P/2$. The power per unit solid angle scattered from the surface, assumed Lambertian, is then

$$J = \frac{P}{2}\frac{\rho\cos\theta}{\pi} \quad \text{watts/steradian,} \tag{19}$$

where θ is an observation angle. We take $\cos\theta \approx 1$, and multiply by the transmission efficiency to find that the power collected from the footprint and impinging on the detector is

$$P_F = \frac{P}{2}\frac{\rho}{\pi}\Omega\eta_{op} = \frac{P}{2\pi}\frac{D_R^2}{R^2}\rho\eta_{op} \quad \text{watts.} \tag{20}$$

The conversion factor to photons is $5\times10^{24}\times\lambda$ photons/joule when λ is expressed in meters, so the photon rate is $5\times10^{24}\times\lambda P_F$. The total number of photons per pixel impinging on the detector in one polarization is this rate multiplied by the dwell time, by the fractional area of the footprint covered by one pixel, and by ½ to account for polarization. Using the synthetic aperture condition $N_{sc} = \lambda R/(2pF)$ from Eq. (18) and $T_{dw} = N_{sc}F/V = \lambda R/(2pV)$, this is

$$\begin{aligned} N_0 &= \frac{5}{2}\times10^{24}\lambda P_F T_{dw}\left(\frac{p}{F}\right)^2 = 5\times10^{24}\lambda\frac{P}{4\pi}\frac{D_R^2}{R^2}\rho\eta_{op}\frac{\lambda R}{2pV}\left(\frac{p}{F}\right)^2 = 5\times10^{24}\frac{P}{8\pi}\frac{\lambda^2}{F^2}\frac{pD_R^2}{RV}\rho\eta_{op} \\ &= 11\left(\frac{P}{1\mathrm{kwt}}\right)\left(\frac{\lambda}{2\,\mu}\right)^2\left(\frac{10\,\mathrm{m}}{F}\right)^2\left(\frac{p}{0.1\mathrm{m}}\right)\left(\frac{D_R}{1\mathrm{m}}\right)^2\left(\frac{1{,}000\,\mathrm{km}}{R}\right)\left(\frac{3.5\,\mathrm{km/sec}}{V}\right)\left(\frac{\rho}{0.1}\right)\left(\frac{\eta_{op}}{0.5}\right) \text{ photons.} \end{aligned} \tag{21}$$

Eqs. (18) and (21), along with the criterion from Eq. (13) that $N_0 = 1/(\eta_d\eta_h)$ provides a near-saturation SNR, contain most of the high-level information needed to decide if a notional design is adequate. The laser required to implement the design must have the capabilities given in Eqs. (15) and (16). We expect to find $1/(\eta_d\eta_h) \approx 2 - 4$ in a well-designed system, so Eq. (21) indicates that the illustrative system is viable. Putting parameters for the Magellan SAR[18] into Eq. (21), P = 50 wt, λ = 12 cm, $F \approx 20$ km, $p \approx 100$ m, $D_R = 3.7$ m, $R \approx 300$ km, $V \approx 10$ km/sec, gives $N_0 \sim 10^7$ photons, which shows why SAR workers don't worry about photon statistics. Eqs. (18) and (21) are design equations, used to determine the hardware parameters (λ, P, D_T, D_R, N_{sc}) needed to produce the desired end-use parameters (p, F, R). We can substitute $\lambda R/F = D_T$ from Eq. (18) into Eq. (21) to find that

$$N_0 = 5\times10^{24}\,\frac{P}{8\pi}\,\frac{pD_T^2D_R^2}{R^3V}\rho\eta_{op} \quad , \tag{22}$$

which shows how the signal scales with range, keeping constant resolution, once the hardware parameters are fixed. Observe that Eq. (22) is independent of λ. The normal radar equation for range dependence of signal is $N_0 \propto R^{-4}$, but Eq. (22) shows that $N_0 \propto R^{-3}$ for SA imaging. The reason is that keeping the same p while doubling R requires doubling the length of the synthetic aperture. With fixed V this takes twice as long, doubling the number of photons that fall on a footprint, thereby reducing the loss of signal with range by one power of R.

Eq. (21) shows the advantage of using the longest wavelength that can give the desired information and/or is technically feasible on a spacecraft: the longest wavelength tends, depending on choices of the other parameters, to produce the largest N_0. The ability to use $N_{sc} > 1$ provides a means of achieving the same footprint and resolution with a longer wavelength by increasing λ and N_{sc} (and D_T) proportionately in Eq. (18). This allows N_0 to be increased and/or laser power to be reduced in Eq. (21), and reduces the PRF given in Eq. (15). An alternative to the baseline design that requires much less laser power is λ = 10 μ, $D_T = D_R$ = 1 m (common transmit and receive aperture, which provides the simplification that K = 1), and N_{sc} = 5. This does incur the relatively mild penalty that the reflectivity of most surface materials tends to be low (~ 5%) in this spectral region, but allows N_0 = 5 with P = 35 wt. Using N_{sc} = 20 and P = 140 wt provides a resolution of p = 2.5 cm, an order of magnitude better than the proposed Mars Reconnaissance Orbiter[19]. If D_R can be larger than one meter, laser power can be further reduced.

The practicality of a visible-light SAL is questionable. Using λ = 0.5 μ and D_T = 0.05 m in Eq. (21) holds out the prospect of p = 2.5 cm from single-look imagery (N_{sc} = 1), but reduces N_0 to 0.2 photons. It would take a 2-m collecting aperture and a 4-kwt laser, or more, to compensate, and a 2-m mirror, diffraction-limited at 0.5 μ, is a formidable challenge indeed for a spacecraft, as is a 4-kwt visible-light laser. The problem can be alleviated by accepting a larger p, but then the resolution of the SA system is probably not sufficiently better than a direct-detection system (with the same D_R) to be worth its additional complication and cost.

The speed of an orbiting platform is different for different bodies, and it may help the reader to know that, for a platform in low orbit around a body of radius R_b, $V \propto \rho^{½}R_b$, where ρ is the body's density, and $\rho^{½}$ varies only from about 1 for small, icy bodies to 2.3 for Mercury. $\rho^{½}$ is 2.0 for Mars, 1.8 for the Moon, between 1.4 and 1.9 for the Gallilean moons of Jupiter

Kyle[10] evaluates a SAL system for the Earth in much the same way as presented in Eq. (21), but does not reduce the result to photons. In our notation, he uses P = 10 wt, λ = 10 μ, p = 0.1 m, $D_T = D_T$ = 0.2 m, R = 200 km, $\rho = \eta_{op}$ = 1, and $V = 8\times10^3$ m/s. Comparing his Eqs. (3) and (4a) shows that his footprint size, denoted D' by him in his Eqs. (2) and (3) and D_I in his Eqs. (10) – (15), is $D' = D_I = F$ = 10 m. Putting these values into Eq. (21) and multiplying by two to include both polarizations yields N_0 = 9 photons, which makes the system viable by our definition (aside from the unrealistic assumptions about reflectivity and transmission efficiency), but falls about two orders of magnitude short of supporting the claim made in the fourth paragraph of his Sec. V that the CNR_{RF} is 331 (Kyle does not consider speckle, so $N_0 = N_p$ and his SNR is our CNR_{RF}).

Kyle's basic error appears to be failing to recognize the discrepancy between (a) the bandwidth needed to match the pulse width of Δt = 0.23 ns stated by him as necessary to give the range resolution specified in his Eq. (19), and (b) the bandwidth from his Eq. (5a) on which the noise expression in his Eq. (16) is based. A wider bandwidth in his Eq. (16) would result in a lower SNR in his Eq. (17). Stated as a time, rather than bandwidth, discrepancy, Kyle's dwell time is $T_{dw} = D_I/V$ = 10/8,000 = 1.25 ms, and he states that there are N^2 pulses in this time, which, with N = 100 (as implied by a footprint of 10 m and a resolution of 0.1 m), implies a pulse time Δt = 0.125 μs, nearly three orders of magnitude greater than 0.23 ns. Kyle does not address this discrepancy. If the 0.125 μs value is substituted into Kyle's Eq. (15), we find D_I = 10 m, as expected; if 0.23 ns is used we find D_I = 5,430 m. Neither value matches Kyle's statement in the sixth paragraph of his Sec. V that D_I = 543 m.

7. Conclusion

This paper argues the theoretical feasibility of an orbit-based SAL at IR wavelengths for achieving centimeter-class resolution on the surfaces of solar system bodies that have little or no atmosphere. SAL's relatively restricted set of observation ranges, a few hundred to a few thousand kilometers, is suited to this problem. At shorter ranges, conventional imaging in visible light can provide high resolution; at longer ranges, excessive laser power and/or real collecting aperture size is required. SAL's limitations of low SNR and area coverage for single-look imagery can be alleviated by multiple images and mosaicking of scenes. A design example has been given that provides 2 - 3 cm resolution on Mars, an order of magnitude better than the Mars Reconnaissance Orbiter[19], currently under study by NASA.

The effect of photon counting statistics on SNR for SAL has been developed. Eqs. (13) and (14) show that a low photon rate imposes a much greater SNR penalty on SAL than on a direct detection system. For SAL, if laser power is not high enough to produce an SNR close to ½ in single-look imagery, then, compared to direct detection, a much larger number of repeated images must be combined to achieve an SNR approaching unity. (But direct detection requires a much bigger real aperture to achieve the same resolution.)

Various engineering difficulties have been touched upon in the course of the discussion. The most obvious are the laser technology issues of developing high-power, space-qualified lasers with fast chirp rates, pulse repetition frequencies of tens or hundreds of kilohertz, and coherence times up to and perhaps beyond tens or even hundreds of milliseconds. Another major engineering problem is providing line-of-sight pointing control consistent with the desired footprint size and capable of executing multiple scans of the scene in order to produce some combination of increased SNR, increased area coverage, increased resolution, and reduced laser power. As stated in Sec. 4, generating the matched filter coefficients for SA signal processing requires compensating for platform vibrations to an accuracy better than the wavelength of the light used, a problem that requires sensitive accelerometers but will be easier to deal with for the smooth motion of a spacecraft than for an airborne system. To an extent, this problem can be handled in post-processing by the focusing methods developed for SAR. Yet another problem is the need to place multiple heterodyne detectors in the receiver focal plane when $D_R > D_T$.

Appendix A – 2-D Gaussian Probability

Following Goodman[15], a complex number, $a\exp(i\theta)$, is called a phasor. Goodman calculates the two-dimensional probability density function that describes the sum of a large number of random phasors. There are two points in this paper to which this PDF is relevant: finding (1) the frequency content of shot noise and the consequent variance with which a detected number of photons is measured by heterodyne detection, and (2) how contributions from the pixels in the beam's ground footprint add up to make the measured signal. In both cases we need to know the sum of N of these random phasors. The sum is an origin-centered 2-D Gaussian distribution described by $\sigma^2 = N\langle a^2\rangle/2$, where $\langle a^2\rangle$ is the expectation value of a^2 over the distribution from which a is chosen, and phase is assumed random and uniformly distributed over $(-\pi, \pi)$. Adding a complex value s, representing a signal, to this distribution displaces its center a distance $|s|$ from the origin, and we may, without loss of generality, take s to be real and non-negative, so the PDF of the sum plus signal is

$$P(x,y) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(x-s)^2+y^2}{2\sigma^2}\right] \quad , \tag{A1}$$

which is taken from Goodman's Eq. (2.9-18) with minor changes in notation. x and y represent the real and imaginary parts, respectively, of a complex number. In extension of the definition of variance for a 1-D Gaussian distribution, the variance of this PDF is

$$\mathrm{Var}\left[P(x,y)\right] \equiv \left\langle (x-s)^2+y^2 \right\rangle = \left\langle (x-s)^2 \right\rangle + \left\langle y^2 \right\rangle = 2\sigma^2 \quad . \tag{A2}$$

When $s = 0$, $(2\sigma^2)^{1/2} = N^{1/2}a_{rms}$ is the rms value of the magnitude of the sum.

The value of σ^2 for heterodyne detection is found by evaluating shot noise. The easiest way to see that shot noise results in white, Gaussian noise in frequency space is to write the current produced in a detector of quantum efficiency η_d by N impinging photons as

$$I(t) = \sum_{n=1}^{\eta_d N} q_e \delta\left(t - t_n\right) \quad , \tag{A3}$$

where q_e is the electronic charge, the sum is over the $\eta_d N$ detected photons, and t_n is the creation time of the nth electron. The Fourier transform of this current is

$$\begin{aligned} \mathrm{FT}\left[I(t)\right] &= \int_0^T \sum_{n=1}^{\eta_d N} q_e \delta\left(t - t_n\right)\exp\left(-2\pi i f t\right) dt \\ &= \sum_{n=1}^{\eta_d N} q_e \exp\left(-2\pi i f t_n\right) \quad . \end{aligned} \tag{A4}$$

Since the t_n are randomly distributed, the second sum in Eq. (A4) is the sum of a large number of phasors with (constant) amplitude q_e and random phase $2\pi f t_n$. Therefore, independently of f, the result is an origin-centered 2-D Gaussian distribution described by Eq. (A1) with $s = 0$ and

$$\sigma^2 = q_e^2 \eta_d N / 2 \quad . \tag{A5}$$

A random number chosen from this distribution is the noise that is added to the DFT component shown in Eq. (4), and, divided by $(q_e\eta_d)^2\eta_h N_L$, appears as the noise term E_m in Table 1.

Eq. (A4) uses the continuous Fourier transform as an easy way to reach the desired result. If the idealized response $\delta(t - t_n)$ is replaced by the actual detector response having finite width, and this width is reasonably densely sampled, the same result is obtained with the discrete Fourier transform used in Eq. (4). The reader who wishes to pursue this topic further may consult Lucke[20] where the properties of photon-limited noise in the DFT of spatial data are explicated at length. The discussion there applies also to the DFT of temporal data, and that paper's Eq. (26) is the equivalent of Eq. (A5) once it is recognized that the total number of photons detected is closely approximated by $\eta_d N_L$ and that the error figure shown in this paper's Fig. 1 is circular (so that, as described in the other paper, $S_{2k} = 0$).

Appendix B – Synthetic Aperture Processing and Resolution

The pixel 0 column of Table 1 shows that the matched filter gives a value of MA_0 for the desired pixel. To justify the claim that the other pixels add to "≈ 0", we first examine the pixel 1 column, which is

$$\begin{aligned}\sum_{m=1}^{M-1} A_1 C_m C_{m-1}^* &= A_1 \sum_{m=1}^{M-1} \exp\left\{ i \frac{2\pi p^2}{\lambda R} \left[\left(m - \frac{M}{2} \right)^2 - \left(m - 1 - \frac{M}{2} \right)^2 \right] \right\} \\ &= A_1 \sum_{m=1}^{M-1} \exp\left\{ i \frac{4\pi p^2}{\lambda R} \left(m - \frac{M+1}{2} \right) \right\}\end{aligned} \tag{B1}$$

where $C_m = \exp\{2\pi i[(m - M/2)p]^2/\lambda R\}$ has been used. The sum on the right side of the second equality is the sum of $M - 1$ unit-amplitude phasors with phase increment $\Delta\phi = 4\pi p^2/\lambda R$. The sum is exactly zero – the phasors "wrap" to zero – if the phase of the last phasor is $2\pi - \Delta\phi$ greater than the phase of the first, *i.e.*, if

$$\frac{4\pi p^2}{\lambda R} \left[\left(M - 1 - \frac{M+1}{2} \right) - \left(1 - \frac{M+1}{2} \right) \right] = \frac{4\pi p^2}{\lambda R} (M - 2) = 2\pi - \frac{4\pi p^2}{\lambda R} \quad , \tag{B2}$$

or

$$\frac{4\pi M p^2}{\lambda R} \left(1 - \frac{1}{M} \right) = 2\pi \quad , \tag{B3}$$

whence

$$p = \frac{\lambda R}{2F} \quad , \tag{B4}$$

where $F = Mp$ is the size of the illuminated footprint and $M >> 1$ has been used. Further, if F is determined by the diffraction-limited resolution of a transmitting aperture with diameter D_T, *i.e.*, $F = \lambda R/D_T$, we find

$$p = \frac{D_T}{2} \quad , \tag{B5}$$

for the resolution of an SA system. Eq. (B5) is the same as, for example, Eq. (1.2.9) of Curlander and McDonough[1].

Eqs. (B1) – (B3) are exact only if pixel 1 consists of a point object at its center. Since the return is actually spread out over the pixel, these equations are approximate, but the basic principle remains: the pixel 1 column of Table 1 makes only a small contribution to the last row because the phasors wrap to (nearly) zero. In the pixel 2 column, the phase increment is twice as big and the wrapping happens faster. The pixel M - 1 column makes a small contribution because it contains only a single term. Intermediate columns make small contributions by a combination of these effects. Finally, all these small contributions are random phasors which add up across the bottom row of the table to give a sum that is small compared to the coherent sum, MA_0, from pixel 0.

References

[1]J. C. Curlander and R. N. McDonough, Synthetic Aperture Radar: Systems and Signal Processing, John Wiley & Sons (New York), 1991.

[2]C. V. Jakowatz, D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach, Kluwer Academic Publishers (Boston), 1996.

[3]R. S. Saunders, A. J. Spear, P. C. Allin, R. S. Austin, A. L. Berman, R. C. Chandlee, J. Clark, A. V. DeCharon, E. M. De Jong, D. G. Griffith, J. M. Gunn, S. Hensley, W. T. K. Johnson, C. E. Kirby, K. S. Leung, D. T. Lyons, G. A. Michaels, J. Miller, R. B. Morris, A. D. Morrison, R. G. Piereson, J. F. Scott, S. J. Shaffer, J. P. Slonski, E. R. Stofan, T. W. Thompson, and S. D. Wall, Magellan Mission Summary, J. Geophys. Res., v. 97, no. E8, 1992, pp. 13,067 – 13,090.

[4]J. W. Goodman, Synthetic Aperture Optics, Sec. 5.4, in Progress in Optics (E. Wolf, ed.), v. VIII, North-Holland (London), 1970, p. 36.

[5]T. J. Green, S. Marcus, and B. D. Colella, Synthetic-aperture-radar imaging with a solid-state laser, Ap. Opt., v. 34, no. 30, 1995, pp. 6,941 – 6,949.

[6]S. Yoshikado and T. Aruga, Short-range verification experiment of a trial one-dimensional synthetic aperture infrared laser radar operated in the 10 μ band, Ap. Opt., v. 39, no. 9, 2000, pp. 1,421 – 1,425.

[7]M. Bashkansky, R. L. Lucke, E. Funk, J. Reintjes, Two-dimensional synthetic aperture imaging in the optical domain, accepted for publication in Optics Letters.

[8]J. H. Shapiro, B. A. Capron, and R. C. Harney, Imaging and target detection with a heterodyne-reception optical radar, Ap. Opt., v. 20, no. 19, 1981, pp. 3,292 – 3,313.

[9]D. Park and J. H. Shapiro, Performance analysis of optical synthetic aperture radars, SPIE Vol. 999 (Richard J. Becherer, ed.), Laser Radar III, 1988, pp. 100 - 116.

[10]T. G. Kyle, High resolution laser imaging system, Ap. Opt., v. 28, no. 13, 1989, pp. 2,651 – 2,656.

[11]C. C. Aleksoff, J. S. Accetta, L. M. Peterson, A. M. Thai, A. Kooster, K. S. Schroeder, R. M. Majewski, J. O. Abshier, and M. Fee, Synthetic aperture imaging with a pulsed CO_2 TEA laser, SPIE Vol. 783 (Richard J. Becherer and Robert C. Harney, eds.), Laser Radar II, 1987, pp. 29 - 40.

[12]R. H. Kingston, Detection of Optical and Infrared Radiation, Springer-Verlag (New York), 1978, p. 27.

[13]P. J. Winzer and W. R. Leeb, Coherent lidar at low signal powers: basic considerations on optical heterodyning, J. Mod. Opt., v. 45, no. 8, 1998, pp. 1,549 – 1,555.

[14]J. H. Shapiro, Target-reflectivity theory for coherent laser radars, Ap. Opt., v. 21, no. 18, 1982, pp. 3,398 – 3,407.

[15]J. W. Goodman, Statistical Optics, John Wiley & Sons (New York), 1985.

[16]J. H. Shapiro and S. S. Wagner, Phase and amplitude uncertainties in heterodyne detection, IEEE J. Quantum Elec., v. QE-20, no. 7, 1984, pp. 803 – 813.

[17]R. Loudon, The Quantum Theory of Light, Clarendon Press (Oxford), 1983, Sec. 4.8.

[18]R. S. Saunders, G. H. Pettengill, R. E. Arvidson, W. L. Sjogren, W. T. K. Johnson, and L. Pieri, The Magellan Venus Radar Mapping Mission, J. Geophys. Res., v. 95, no. B6, 1990, pp. 8,339 – 8,355.

[19]J. Graf and R. Zurek, The Mars Reconnaissance Orbiter Mission, presented as paper IAF-01-Q.3.a.07, at the 52nd International Astronautical Congress at Toulouse, France, Oct. 1 - 5, 2001 (no separate proceedings published). Available from the Linda Hall Library of Science, Engineering and Technology, Reference Department, 5109 Cherry Street, Kansas City, MO 64110 (816-363-4600). (American Institute of Aeronatics and Astronautics accession number #A01-41996).

[20]R. L. Lucke, Fourier-space properties of photon-limited noise in focal plane array data, calculated with the discrete Fourier transform, J. Opt. Soc. Am. A, v. 18, no. 4, 2001, pp. 777 – 790.

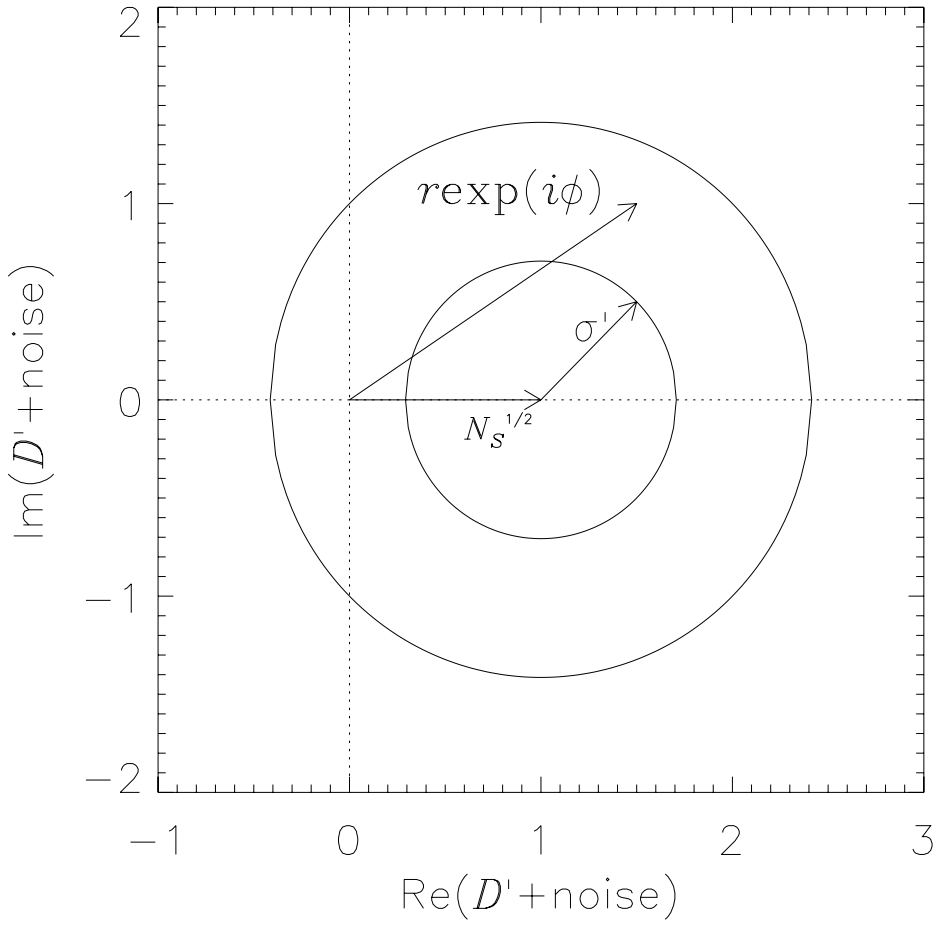

Figure 1. Distribution of measured values with $N_S^{½} = 1$, $\phi_S = 0$, for the ideal case $\sigma'^2 = ½$ (*i.e.*, $\eta_d = \eta_h = 1$), showing 1- and 2-sigma contours. For the non-ideal case, σ'^2 is increased in accordance with Eq. (5). $r\exp(i\phi)$ represents a particular measurement taken from this distribution. $r = n^{½}$, where n is the number of photons detected in the measurement (see text).

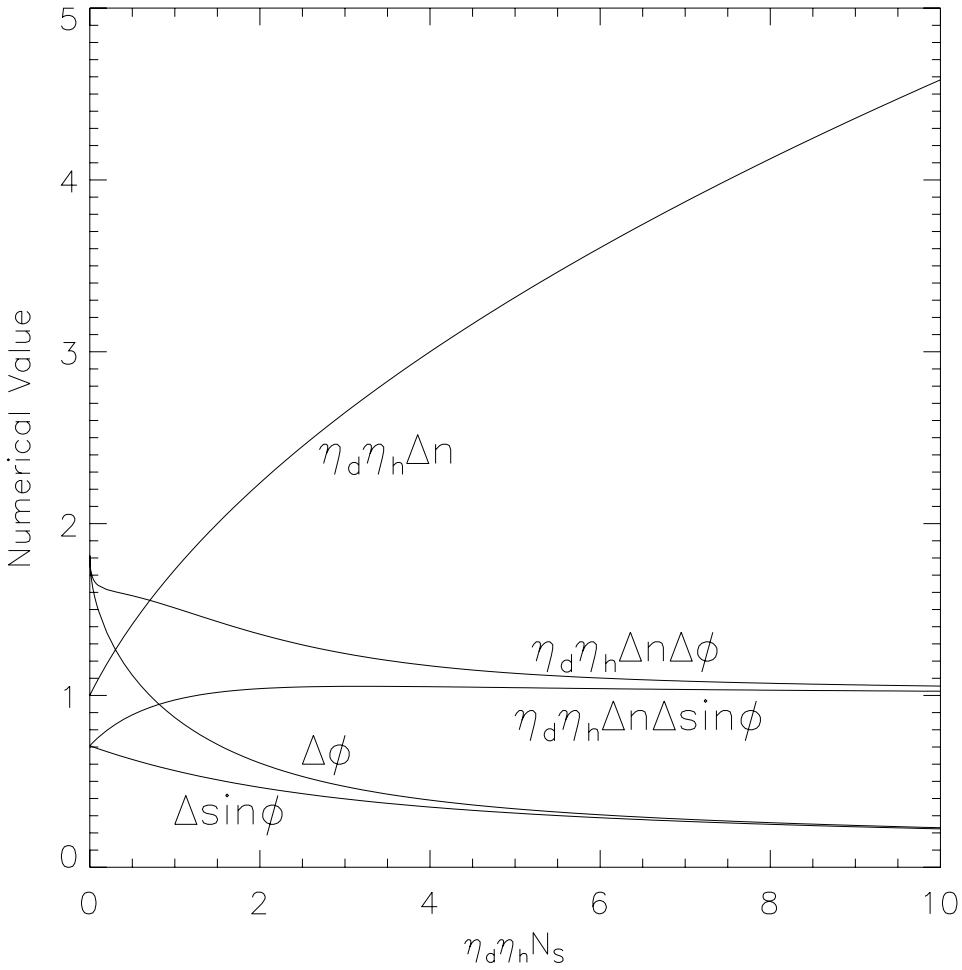

Figure 2. Number and phase uncertainties for phase-sensitive heterodyne detection. $\eta_d\eta_h\Delta n = \sqrt{2\eta_d\eta_h N_S + 1}$. For $N_S = 0$, $\Delta\phi = \pi/\sqrt{3}$ and $\Delta\sin\phi = 1/\sqrt{2}$.

Figure 3. Side view of transmitted wave front encountering one range resolution element. Azimuthal resolution elements (pixels) are labeled 0 to *M*-1 (see text). The transmitter is traveling to the left.